Calhoun: The NPS Institutional Archive

DSpace Repository

Theses and Dissertations	1. Thesis and Dissertation Collection, all items

1981-03

The one-sample K-S test: TI-59 programs for discrete and continuous random variables

Brown, Douglas Warren; Barr, Donald R.

Monterey, California. Naval Postgraduate School

http://hdl.handle.net/10945/20627

Downloaded from NPS Archive: Calhoun

http://www.nps.edu/library

Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

THE ONE-SAMPLE K-S TEST: TI-59 PROGRAMS FOR DISCRETE AND CONTINUOUS RANDOM VARIABLES

Douglas Warren Brown

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

THE ONE-SAMPLE K-S TEST: TI-59 PROGRAMS
FOR DISCRETE AND CONTINUOUS RANDOM VARIABLES

by

Douglas Warren Brown

March 1981

Thesis Advisor: D.R. Barr

Approved for public release, distribution unlimited

T199251

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): The One-Sample K-S Test: TI-59 Programs for Discrete and Continuous Random Variables	5. TYPE OF REPORT & PERIOD COVERED: Master's Thesis, March 1981
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Douglas Warren Brown	8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Postgraduate School, Monterey, California 93940	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Naval Postgraduate School, Monterey, California 93940	12. REPORT DATE: March 1981
	13. NUMBER OF PAGES: 72
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report)
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Kolmogorov- Smirnov, K-S, TI-59, hidden registers, pending arithmetic registers, Discrete K-S, Continuous K-S

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The Kolmogorov-Smirnov (K-S) goodness-of-fit test is a nonparametric test if the random varible is continuous. If the random variable is discrete the K-S test is not nonparametric. Two programs for the TI-59 programmable calculator are presented, "Continuous K-S" and "Discrete K-S". They are designed to conduct a K-S test and display the significance level. User instructions and several sample

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

problems are presented. Use of the "hidden registers" in the TI-59 is discussed. The hidden registers are used extensively in the program Discrete K-S.

Approved for public release, distribution unlimited

The One-Sample K-S Test: TI-59 Programs for Discrete and Continuous Random Variables

by

Douglas Warren Brown
Lieutenant, Supply Corps, United States Navy
A.B., Whitman College, 1974

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 1981

ABSTRACT

The Kolmogorov-Smirnov (K-S) goodness-of-fit test is a nonparametric test if the random variable is continuous. If the random variable is discrete the K-S test is not nonparametric.

Two programs for the TI-59 programmable calculator are presented, "Continuous K-S" and "Discrete K-S". They are designed to conduct a K-S test and display the significance level. User instructions and several sample problems are presented.

Use of the "hidden registers" in the TI-59 is discussed. The hidden registers are used extensively in the program Discrete K-S.

TABLE OF CONTENTS

I. INTRODUCTION ---------- 6

II. THE KOLMOGOROV-SMIRNOV (K-S) TEST ---------- 7

III. TRANSFORMATIONS OF SAMPLE DATA ---------- 15

IV. THE K-S TEST: CONTINUOUS RANDOM VARIABLE ------- 17

A. DISTRIBUTIONS OF THE TEST STATISTICS -------- 17

B. USER INSTRUCTIONS FOR "CONTINUOUS K-S" ------ 18

C. USER INSTRUCTIONS TO EVALUATE A GIVEN TEST STATISTIC ---------- 24

V. THE K-S TEST: DISCRETE RANDOM VARIABLE --------- 25

A. DISTRIBUTIONS OF THE TEST STATISTICS -------- 25

B. USER INSTUCTIONS FOR "DISCRETE K-S" --------- 27

APPENDIX A CONTINUOUS K-S: ENTRY PROCEDURE, PROGRAM LISTING, AND EXAMPLES ---------- 32

APPENDIX B DISCRETE K-S: ENTRY PROCEDURE, PROGRAM LISTING, AND EXAMPLES ---------- 50

APPENDIX C CONDENSED USER INSTUCTIONS FOR CONTINUOUS K-S AND DISCRETE K-S ---------- 65

APPENDIX D USE OF PENDING ARITHMETIC REGISTERS IN THE TI-59 ---------- 68

LIST OF REFERENCES ---------- 71

INITIAL DISTRIBUTION LIST ---------- 72

I. INTRODUCTION

The subject of this thesis is the one-sample Kolmogorov-Smirnov (K-S) goodness-of-fit test. The purpose of the thesis is to provide the reader with programs enabling him to perform K-S tests on the TI-59 programmable calculator.

The distinction between continuous and discrete random variables is a central issue. Different computational procedures are appropriate when testing a discrete random variable than when testing a continuous random variable. Two separate TI-59 programs are provided: "Continuous K-S" and "Discrete K-S".

Section II is a general discussion of the K-S test. Section III explains the technique of transforming sample data before performing a K-S test. Section IV discusses the distribution of test statistics from a continuous K-S test and provides detailed instructions on how to use the Continuous K-S program. Section V is an analagous discussion for a discrete K-S test.

Appendicies A and B are entry instuctions, program listings, and example problems for Continuous K-S and Discrete K-S, respectively. Appendix C presents condensed user instructions for both programs. Appendix D discusses the use of the "hidden registers" in the TI-59.

II. THE KOLMOGOROV-SMIRNOV (K-S) TEST

A. OVERVIEW OF THE K-S TEST

Statisticians are often faced with observations of a random variable based on a population with an unknown probability distribution. A goodness-of-fit test may be used to test certain hypotheses about the relationship between the population's distribution and some specified distribution. The one-sample Kolmogorov-Smirnov (K-S) test is a goodness-of-fit test. The term "one-sample" is used because a "two-sample" K-S test has also been developed. The two-sample test examines the relationship between two unknown probability distributions on the basis of two random samples from different populations. Only the one-sample K-S test is discussed herein. It will be refered to below as simply the "K-S test".

The K-S test is conducted by comparing S_n, the empirical cumulative distribution function (CDF), defined below, with H, the hypothesized CDF. If the discrepancy between S_n and H, called the "test statistic", is large then we have strong evidence that F, the population CDF, does not equal H. The probability of observing a test statistic at least as large as that realized is called the "significance level" of the test.

A K-S test is only valid if outcomes on the random variable lie on an ordinal scale. If a random variable maps events into cells between which no ordinal relationship exists then the chi square goodness-of-test is recommended [Ref. 1: pp. 237].

Any goodness-of fit test can be either a hypothesis test or a significance test. For a hypothesis test we specify a null hypothesis, H_o, an alternative hypothesis, H_1, and an "alpha level". After we compute the significance level we compare it with the alpha level. If the significance level is less than or equal to the alpha level then we reject H_o and accept H_1. Otherwise, we do not reject H_o. To conduct a significance test we simply compute the significance level; we are not formally concerned about rejecting H_o.

The requisite computational techniques, yet to be discussed, are identical for hypothesis and significance testing. For either type of test the test statistic must be computed and then evaluated by computing the significance level. Hence, the discussion of the K-S test below will focus on those two steps.

Let X_1 , X_2, . . . , X_n be independent observations on the random variable X which has CDF F. Assume an ordinal relationship exists between all possible values of X. If $X_{(1)}$, $X_{(2)}$, . . . $X_{(n)}$ represent the n observations arranged in ascending order then the empirical CDF, S_n is defined as:

$$S_n(x) = \begin{cases} 0 & \text{if} \quad x < X_{(1)} \\ k/n & \text{if} \quad X_{(k)} \leq x < X_{(k+1)}, \quad k = 1, 2, \ldots, n-1 \\ 1 & \text{if} \quad x \geq X_{(n)} \end{cases}$$

The K-S test may be used to test the following three pairs of hypotheses given a specified CDF H. The appropriate test statistic is also listed.

1. H_o: $F(x) \geq H(x)$ for all x

   H_1: $F(x) < H(x)$ for some x

   Test statistic: $D^- = \sup_x(F(x) - S_n(x))$

2. H_o: $F(x) \leq H(x)$ for all x

   H_1: $F(x) > H(x)$ for some x

   Test statistic: $D^+ = \sup_x(S_n(x) - H(x))$

3. H_o: $F(x) = H(x)$ for all x

   H_1: $F(x) \neq H(x)$ for some x .

   Test statistic: $D = \max(D^-, D^+)$

Note that by simply specifying H_o we implicitly specify H_1 since H_1 is the negation of H_o. It should also be noted that the first two null hypotheses listed pertain to what is commonly called "one-sided" tests while the third pertains to a "two-sided" test. The test statistics, D^-, D^+, and D, are collectively called the "K-S statistics".

B. CALCULATING THE VALUE OF THE TEST STATISTIC

1. Discrete Random Variable

When examining a discrete random variable the the test statistics D^- and D^+ are found by comparing the

empirical and hypothesized CDFs at each possible value of the random variable. More precisely,

$$D^+ = \max \{0, \max[S_n(x) - H(x)]\}$$

and $$D^- = \max \{0, \max[H(x) - S_n(x)]\} .$$

Figure 1 illustrates this approach.

2. Continuous Random Variable

If we are examining a continuous random variable we calculate D^+ and D^- using the values of the empirical and hypothesized CDFs at the observed values in our sample. The computation of D^+ is analagous to the procedure used with a discrete random variable. The computation of D^-, however, is conceptually different. Using previous notation:

$$D^+ = \max\{0, \max[S_n(X_{(i)}) - H(X_{(i)})]\}, \quad i = 1, 2, \ldots, n$$

$$D^- = \max\{0, \max[H(X_{(i)}) - S_n(X_{(i-1)})]\}, \quad i = 1, 2, \ldots, n$$

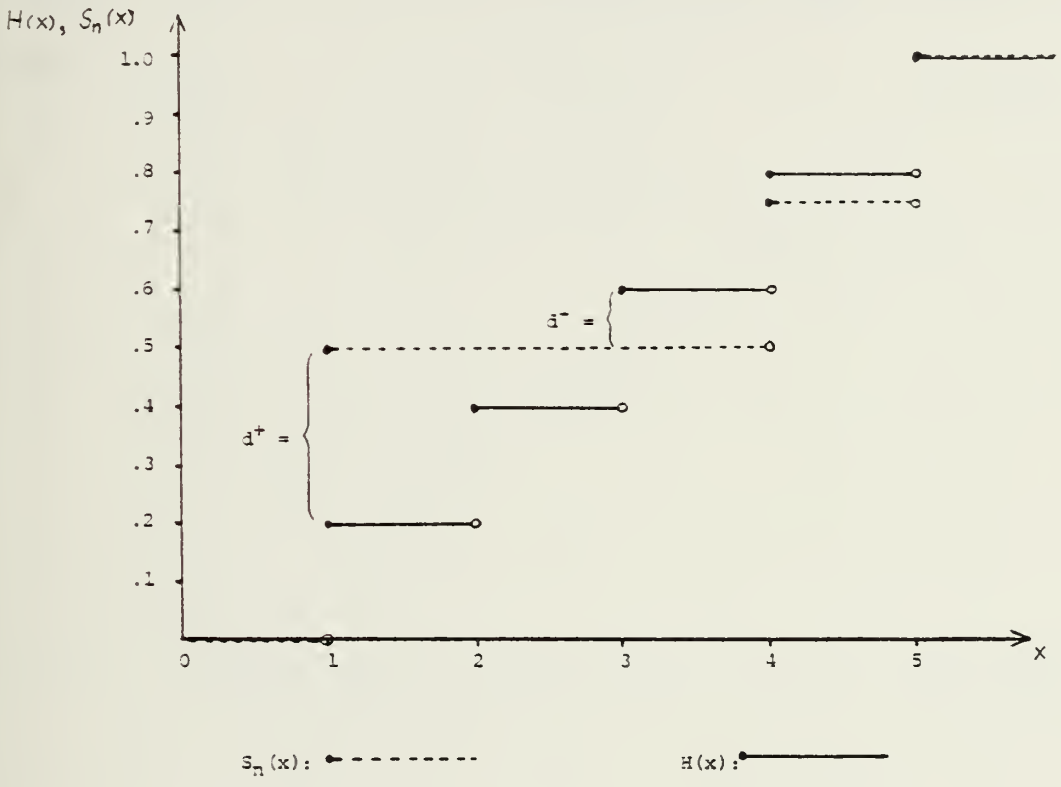
We define $X_{(0)} = -\infty$, thus $S_n(X_{(0)}) = 0$. Figure 2 illustrates this approach.

C. HISTORICAL ORIGINS OF THE K-S TEST

The K-S test was first introduced by Kolmogorov in 1933 [Ref. 2] and was further developed by Smirnov [Ref. 3]. The essence of the original work in this field was the derivation of the distribution of the K-S statistics D^+, D^-, and D. Continuity of the random variable was a key assumption in the development of the test. The statistics D^+ and D^- were discovered to follow the same distribution and this distribution was derived exactly. The distribution

Hypothesized distribution: $H(x) = x/5$, $x = 1, 2, 3, 4, 5$

Sample data: 1, 1, 4, 5 (sample size n = 4)



$$d^+ = \max\{0, \max_x[S_n(x) - H(x)]\}, \quad x = 1, 2, 3, 4, 5$$
$$= \max\{0, \max[(.5-.2), (.5-.4), (.5-.6), (.75-.8), (1-1)]\}$$
$$= \max\{0, \max[.3, .1, -.1, -.05, 0]\}$$
$$= .3$$

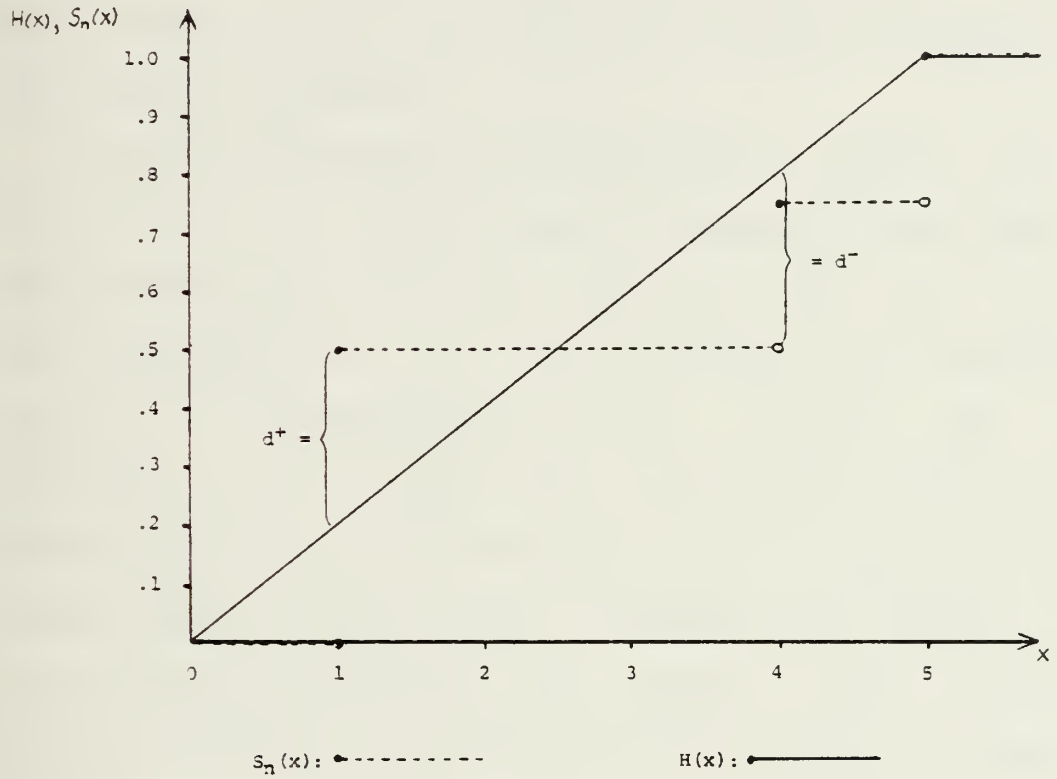
$$d^- = \max\{0, \max_x[H(x) - S_n(x)]\}, \quad x = 1, 2, 3, 4, 5$$
$$= \max\{0, \max[(.2-.5), (.4-.5), (.6-.5), (.8-.75), (1-1)]\}$$
$$= \max\{0, \max[-.3, -.1, .1, .05, 0]\}$$
$$= .1$$

Figure 1

Example: Test Statistics from a Discrete K-S

Hypothesized distribution: $H(x) = \begin{cases} 0 & , \quad x < 0 \\ x/5 & , \quad 0 \leq x \leq 5 \\ 1 & , \quad x > 5 \end{cases}$

Sample data: 1, 1, 4, 5 (ties occur from rounding error)



$$d^+ = \max \{0, \max_i [S_n(x_{(i)}) - H(x_{(i)})]\} \ , \ i = 1, 2, 3, 4$$
$$= \max\{0, \max[(.5-.2), (.5-.2), (.75-.8), (1-1)]\}$$
$$= \max\{0, \max[.3, .3, -.05, 0]\}$$
$$= .3$$

$$d^- = \max\{0, \max_i [H(x_{(i)}) - S(x_{(i-1)})]\} \ , \ i = 1, 2, 3, 4$$
$$= \max\{0, \max[(.2-0), (.2-.5), (.8-.5), (1-.75)]\}$$
$$= \max\{0, \max[.2, -.3, .3, .25]\}$$
$$= .3$$

note: by definition, $x_{(\phi)} = -\infty$, thus, $S_n(x_{(\phi)}) = 0$

Figure 2

Example: Test Statistics from a Continuous K-S

of D was closely approximated. The accuracy of the approximation was shown to be very nigh when the significance level of a test is small. Bradley [Ref. 4: pp. 296] has presented a derivation of the distribution of the K-S statistics which is based upon, but more readable than, the original arguments.

When the hypotheszed distribution, H, equals the true distribution, F, and the random variable is continuous then the distributions of the K-S statistics have been shown to be independent of H. For this reason the continuous K-S test is a "nonparametric" test. Tables are widely available, e.g., Bradley [Ref. 4: pp. 367], delineating the distribution of the K-S statistics as a function only of sample size. This is done by showing the "critical values" of the test statistics for certain widely used alpha levels. The user is thus relieved of having to compute the significance level when performing a hypothesis test. He need only compute the test statistic and compare it to the critical value.

It is incorrect to consider the K-S test to be a nonparametric test when the random variable of interest is discrete. Conover [Ref. 5] has shown how the distributions of the K-S statistics for a discrete random variable are dependent upon both the particular hypothesized distribution and the sample size. If the K-S statistics which result from testing a discrete random variable are assumed to follow the

distributions derived by Kolmogorov then the significance level computed will be larger than the true significance level. Similarly, reliance on tables such as Bradley's will impose critical values that are unnecessarily large for a given alpha level. By either relying directly upon Kolmogorov's distributions or upon tables such as Bradley's the effect is the same. The evidence against H_o, the null hypothesis, will appear much weaker (or "less significant") than it really is.

III. TRANSFORMATION OF SAMPLE DATA

The K-S test is usually applied without employing transformations on the sample data. However, since it is sometimes convenient to transform data before conducting a K-S test, the Continuous K-S program is desgned to utilize transformations as a user option. Transformations are not recommended for a discrete K-S test.

Assume W is a continuous random variable with CDF H and a transformation M is defined such that M(W) has a known CDF, G. Further assume that we have observations on a continuous random variable X with some unknown CDF, F. We may conduct a K-S test of our hypothesis concerning the relationship between F and H in two ways. Without using any transformation the empirical CDF S_n may be compared to H. Alternatively, the empirical CDF S_n' may be compared to G, where S_n' is based upon observations of the random variable M(X). Both methods will yield the same test statistic and significance level.

A frequent use of this technique is to define the transformation M to be the function H, i.e., to invoke the probability integral transformation [Ref. 6: pp. 202]. For instance, given a continuous random variable X with CDF H the probability integral transformation tells us that the random variable H(X) is distributed uniform (0,1). If X is

the population random variable in a K-S test and H is the hypothesized CDF, we may wish to compare the empirical CDF S_n' from observations on H(X) to a uniform (0,1) CDF. This might be easier than directly comparing S_n to H. Example 3 in Appendix A illustrates this approach.

As another example, it is sometimes convenient to employ a transformation when the hypothesized distribution is lognormal. By definition, the random variable W is distributed lognormal (μ, σ^2) if and only if ln(W) is distributed normal (μ, σ^2) [Ref. 6: pp. 117]. Thus, if we have observations on a random variable X and we hypothesize that X is distributed lognormal (μ, σ^2), we may simplify computation by transforming the data. By treating the values of ln(x) as our sample data and normal (μ, σ^2) as our hypothesized distribution we can take advantage of the normal CDF subroutine in the TI-59 calculator. See example 4 in Appendix A.

IV. THE K-S TEST: CONTINUOUS RANDOM VARIABLE

A. DISTRIBUTIONS OF THE TEST STATISTICS

The test statistic in a K-S test is a random variable with a particular probability distribution. After conducting a K-S test with test statistic T ($T = D^+$, D^-, or D) we observe some value, t, from the distribution of T. To find the test's significance level we compute $P(T \geq t)$. If $T = D^+$ or D^- it was shown by Birnbaum and Tingey [Ref. 7] that

$$(IV.1) \qquad P(T \geq t) = t \sum_{j=\emptyset}^{[n(1-t)]} \binom{n}{j} (1-t-j/n)^{n-j} (t+j/n)^{j-1}$$

where n is sample size, $\binom{n}{j} = n!/j!(n-j)!$, and $[n(1-t)]$ is the largest integer contained in $n(1-t)$. For $T = D$ the probability value above should be doubled. The significance levels computed with equation IV.1 will be exact for one-sided tests. For a two-sided test the doubled result will be an upper bound on the true significance level. The difference between the upper bound and the exact value will be negligible when $P(D \geq t)$ is small.

For large sample sizes there are useful approximations to equation IV.1, but their accuracies are marginal for sample sizes as small as $n = 50$. On the other hand, the

difficulty in using equation IV.1 for hand computation is considerable for sample sizes greater then n = 10. Furthermore, even finding the value of the test statistic can be tedious. The Continuous K-S program implements equation IV.1 using a TI-59 calculator and is listed in Appendix A along with several example problems. The program is designed for sample sizes less than n = 55 and is used as follows.

B. USER INTRUCTIONS FOR "CONTINUOUS K-S" PROGRAM

STEP ONE - Specifying the test: The user must choose which continuous probability distribution, H, will be compared against the empirical distribution from the sample data (or the transformed sample data). The program's options for H are: normal, exponential, Weibull, or uniform. If H is normal then the Applied Statistics module must be used, otherwise, any module may be used in the calulator. The user must specify the appropriate parameter(s) of H.

STEP TWO - Entering the data: Once the test is specified, read both sides of Card 1 (Continuous K-S) into the TI-59. A printer may be used but it is not required. To enter the data (in any order):

Enter	Press	Display	Printer (Optional)
	E	399.69	
X_i ($i = 1, \ldots, n$)	A	i	X_i
	C	1	"INPUT DATA"

sorted data	$X_{(1)}$	Ø1
	$X_{(2)}$	Ø2
	.	.
	.	.
	.	.
	$X_{(n)}$	n
ignore	Ø	n+1
	Ø	n+2
	.	.
	.	.
	.	.

There will be a delay after pressing "C" while the calculator sorts the data into ascending order. For a sample of size 1Ø this will take about 1 minute while for the maximum sample size of 54 the sorting requires about 15 minutes. The user may find it preferable to sort the data himself. If the user enters the data in ascending order the delay will be between 1Ø and 75 seconds after pressing "C", depending upon sample size. If the printer is used a listing of the ordered input data will be produced.

If the sorted data are to be tested directly without the use of a transformation the next step should be omitted.

STEP THREE - Transforming the data: If a transormation M is desired the user must provide it as a subroutine for Card 1. The subroutine must begin with "LBL E′" and end with "INV SBR". After Card 1 has been read into the

calculator the user should enter the subroutine, beginning at program step 325. It must not exceed 75 steps in length. For a sample size of size n, the subroutine must use only registers numbered n+1 through 57. The subroutine should be written to be called when value X is currently in the "x-register" and M(X), i.e., the transformed value, is to be returned to the main program by the subroutine.

For example, assume a transformation t defined as :

$$M(X) = (X-2)/3 .$$

After reading Card 1 into the calculator the appropriate steps are:

Press	Display
GTO 325	
LRN	325 ØØ
LBL	326 ØØ
E′	327 ØØ
−	328 ØØ
2	329 ØØ
)	330 ØØ
÷	331 ØØ
3	332 ØØ
)	333 ØØ
INV SBR	334 ØØ
LRN	

After entering the tranformation the user should test it before proceeding. In the above example M(8) = 2. To test: enter "8", press "E′", and the calculator should display "2".

After validating the transformation subroutine the user should press "B′". When then transformation is completed $M(X_{(n)})$ is displayed. If a printer is used the transformed

data are listed. The user is now ready to perform the K-S test.

If the hypothesized distribution, H, is lognormal (μ, σ^2) then a different procedure is followed since the appropriate transformation is provided as an optional subroutine on Card 1. After entering the data and pressing "C" the program will sort (or confirm that the user has sorted) the original data. The user then presses "A'". This will transform each observed value of X to ln(X). The calculator will display $\ln(X_{(n)})$ a few moments later and, if a printer is used, a listing of the transormed data will be produced. The K-S test can then be conducted by comparing the transformed data to a normal (μ, σ^2) distribution.

If H is lognormal then a K-S test can be performed with Continuous K-S only if the sample data is transformed. A lognormal random variable is strictly positive. Hence, if any input values are nonpositive the calculator will flash after "A'" is pressed, indicating an error.

STEP FOUR - Conducting the test: Both sides of Card 2 (Continuous K-S) should be read into the calculator. Card 2 is initialized differently depending on the hypothesis being tested.

a. Initialization:

(1) press A if H_0: $F(x) = H(x)$ for all x
H_1: $F(x) \neq H(x)$ for some x

(2) press B if H_0: $F(x) \geq H(x)$ for all x
H_1: $F(x) < H(x)$ for some x

(3) press C if H_0: $F(x) \leq H(x)$ for all x
H_1: $F(x) > H(x)$ for some x

In all cases "Ø" is displayed. If the user changes his mind about the direction of his hypotheses after initializing he should press "RST" before pressing the correct user label.

b. Defining H and computing significance level (S.L.):

(1) Normal (μ, σ^2)

Enter	Press	Display
μ	A′	μ
σ	R/S	S.L.

(2) Exponential (λ) $\{ H(x) = 1 - e^{-\lambda x} ; x \geq 0 \}$

Enter	Press	Display
λ	B′	1
1	R/S	S.L.

(3) Weibull (λ, α) $\{ H(x) = 1 - e^{-(\lambda x)^{\alpha}} ; x \geq 0 \}$

Enter	Press	Display
λ	B′	1
α	R/S	S.L.

(4) Uniform (a,b)

Enter	Press	Display
a	C′	a
b	R/S	S.L.

The computation time will vary from 1 minute with sample size of 10 up to 20 minutes for the maximum sample size of 54. After the computation is finished and the significance level is displayed the user may recall the test statistic from register ØØ.

STEP FIVE - Evaluating the significance level:

The significance level computed may be considered exact if conducting a one-sided test. After conducting a two-sided test if the significance level computed is less than .1 it may be considered accurate to 4 decimal places. [Ref. 4: pp. 301] Since most alpha levels are at most .1 it is anticipated that virtually all user requirements will be satisfied with this degree of accuracy.

When the data "agree closely" with the hypothesized distribution then the significance level computed in a two-sided test may be considerably larger than the true value; it may even exceed 1. When this happens the user should conclude that the sample data is consistent with the null hypothesis.

C. USER INSTRUCTIONS TO EVALUATE A GIVEN TEST STATISTIC

If a user already knows the value of the test statistic from a K-S test on a continous random variable he may use Continous K-S in the TI-59 to evaluate the significance level. This will provide the user with more precise information than he could get by consulting a table of critical values.

Assume that a K-S test was conducted with sample size n and that the observed value of the test statistic was t. The user should first repartition the calculator by entering "7" and pressing "Op 17". The calculator will display "399.69". Both sides of Card 2 should then be read into the calculator. The significance level (S.L.) of the test may be computed by the following steps:

Enter	Press	Display
n	STO 59	n
t	STO ØØ	t
	E′	S.I.

If the test was one-sided the significance level displayed may be considered exact. If n and t were from a two-sided test the significance level displayed should be doubled. The comments in STEP FIVE of the USER INSTRUCTIONS above apply when using the program to evaluate a known test statistic from a two-sided test.

V. THE K-S TEST: DISCRETE RANDOM VARIABLE

A. THE DISTRIBUTIONS OF TEST STATISTICS

The distributions of the K-S statistics are not independent of the hypothesized distribution, H, when the random variable is discrete. If a test statistic in such a case is evaluated using formula IV. 1 (employed by Continuous K-S) the significance level computed should be viewed as a very conservative upper bound on the true significance level. Conover [Ref. 5] has derived the distribution of the K-S statistics when H is discrete. He has shown that

$$(V.1) \qquad P(D^{+} \geq t) = \sum_{j=\emptyset}^{[n(1-t)]} \binom{n}{j} f_j^{n-j} e_j$$

with notation defined as:

1. $[n(1-t)]$ is the largest interger in $n(1-d)$
2. $\binom{n}{j} = n!/j!(n-j)!$
3. $f_j := P\{X < H^{-1}(1 - t - j/n + eps)\}$

   where X has distribution function H and

$$H^{-1}(p) = \begin{cases} +\infty & , \quad p > 1 \\ \sup\{x;\ H(x) < p\} & , \quad \emptyset < p < 1 \\ -\infty & , \quad p < \emptyset \end{cases}$$

4. $e_0 = 1$ and

$$e_k = 1 - \sum_{j=0}^{k-1} \binom{k}{j} f_j^{k-1} e_j , \quad k \geq 1$$

5. The value "eps" is a positive number that approaches, but does not reach, zero. In the notation of the limit arguments from calculus: $eps \rightarrow 0^+$.

Similarly,

$$(V.2) \qquad P(D^- \geq t) = \sum_{j=0}^{[n(1-t)]} \binom{n}{j} c_j^{n-j} b_j$$

with additional notation defined as:

1. $c_j = P\{X \geq H(t + j/n)\}$
2. $b_0 = 1$ and

$$b_k = 1 - \sum_{j=0}^{k-1} \binom{k}{j} c_j^{k-1} b_j , \quad k \geq 1 .$$

Finally, for a two-sided test, the distribution of D is closely approximated as follows for small values of t:

(V.3.) $P(D \geq t) = P(D^+ \geq t) + P(D^- \geq t) - P(D^+ \geq t)P(D^- \geq t)/2$

with a maximum error term of $P(D^+ \geq t)P(D^- \geq t)/2$. For significance levels less than or equal to .1 the error will be neglible.

B. USER INSTRUCTIONS FOR "DISCRETE K-S" PROGRAM

The equations defined above require an unacceptable amount of effort to employ by hand computation. The Discrete K-S program implements them on the TI-59 calculator and is listed in Appendix B along with several example problems.

STEP ONE - Checking the Problem Size:

Discrete K-S is designed to test a discrete random variable with m possible outcomes. It may be necessary to aggregate possible outcomes into cells since m must be less than 21. With a sample of size n the program can perform a K-S test if $\{[n(1-t)] + m\} \leq 54$. By convention, $[n(1-t)]$ is defined to be the largest integer in $n(1-t)$. Of course, the value of t, the test statistic, is unknown before the test is conducted but if $(n + m) \leq 54$ then the requirement above is certainly met. If $(n + m)$ is only slightly greater than 54 the data should be entered and processed with Card 1 as explained below. The program will compute the value of the test statistic and indicate if the problem is too large to complete.

STEP TWO - Defining the Hypothesized Distribution:

Each possible value of the random variable being investigated must be classified as falling in one of the m cells. The user must specify the hypothesized distribution, H, in terms of the cells. This may be done in one of two

ways. For each cell i the user can specify either n_i or H_i. For a random observation, X, on the random variable we define h_i and H_i as follows:

$$h_i = P(X \text{ is in cell } i) , \qquad i = 1, 2, \ldots, m$$

$$H_i = P(X \text{ is in cell } j \; ; \; j \leq i) , \qquad i = 1, 2, \ldots, m$$

We call h_i the mass value of cell i and H_i the cumulative mass value of cell i. The number of the n sample observations that fall into cell i is denoted X_i.

STEP THREE - Initializing the Program and Entering the Data:

Read both sides of Card 1 (Discrete K-S) into the calculator and proceed as follows.

Enter	Press	Display	
	E (if H defined by h_is)	Ø	
	E′ (if H defined by H_is)	Ø	
m	R/S	m	
n	R/S	1	
h_1 or H_1	R/S	2	
h_2 or H_2	R/S	3	
.	.	.	
.	.	.	
.	.	.	
h_{m-1} or H_{m-1}	R/S	m	
h_m or H_m	R/S	1	(see Note 1)
X_1	R/S	2	
X_2	R/S	3	
.	.	.	
.	.	.	
.	.	.	
X_{m-1}	R/S	m	
X_m	R/S	Ø	(see Note 2)

Note 1: If H was entered by mass values n_i , i = 1, 2, . . ., m and ($h_1 + h_2 + . . . + h_m$) is not between .99 and 1.01 then error code "22" will be diplayed after n_m is entered. The user can press "x:t" to display ($h_1 + h_2 + . . . + h_m$). Ideally, the mass values will always sum to 1 but the user may be using rounded values which do not. If the user wishes to proceed with the K-S test after error code "22" is displayed he should press "A'". This will cause the mass values to be normalized so they will sum to 1. Normally, this should only be done if a more accurate statement of mass values is unavailable.

Note 2: If ($X_1 + X_2 + . . . + X_m$) ≠ n then error code "33" will be displayed. The test cannot proceed.

STEP FOUR - Computing the Test Statistic:

After completing the Initialization and Data Entry sequence the test statistic, t, should be computed.

If H_o: $F(x) = H(x)$ then press A.

If H_o: $F(x) \geq H(x)$ then press B.

If H_o: $F(x) \leq H(x)$ then press C.

After processing the data and computing the appropriate test statistic, t, the calculator will either display t or "44". The latter indicates that the problem is too large for Discrete K-S. The user can recall t from register ØØ. If t is displayed (always a number between Ø and 1) then the user can proceed with the test.

STEP FIVE - Computing the Significance Level:

After the test statistic is displayed the program has automatically repartitioned the calculator to allow Card 2 to be read into the calculator. The user is cautioned not to press "CP" or "RST" between using Card 1 and Card 2. This is important because certain flags are set in the calculator during the previous steps.

To compute the significance level of the test simply read Card 2, both sides, into the calculator and press "A". If a one-sided test is conducted, then the significance level displayed may be considered exact. If a two-sided test is conducted then the calculator displays an approximate significance level. The following steps may be used to judge the quality of the approximation:

Press	Display
B	upper bound on S.L.
B, x≷t	lower bound on S.L.
C	maximum error in point estimate of S.L.
D	recall point estimate of S.L.

After the test is completed the calculator can be restored to normal partitioning by pressing "E".

The user should be aware that the time required to compute a significance level can be considerable. The following table may be used to estimate computing time.

Problem Size:		Time to compute significance level:	
n	m	one-sided test	two-sided test
10	5	3 min, 30 sec	7 min
20	10	45 min	90 min
40	12	2 hours	4 hours

If the user needs the capability to handle larger problems and/or process data quickly a FORTRAN program has been developed by Allen [Ref. 9] for this purpose.

APPENDIX A

CONTINUOUS K-S: ENTRY PROCEDURE, LISTING, AND EXAMPLES

A. ENTRY PROCEDURE

The program is organized in two distinct parts, Card 1 and Card 2. Card 1 should be recorded on a magnetic card using banks 1 and 2 [Ref. 8: pp. VII-2] with the calculator in standard partitioning [Ref. 8: pp. V-22]. Card 2 should also be recorded on banks 1 and 2 but with partitioning 399.69.

It should be noted in the program listing below that Cards 1 and 2 both use the "Dsz" command followed by register numbers greater than 9. The user's manual [Ref. 8: pp. V-63] implies that this will not work but, in fact, it does. The user has to use a special entry technique to get the program steps into the calculator correctly, however. One very convenient way to accomplish this is illustrated easily by an example.

Assume that the steps at one point in the program are "Dsz, 58". The user can first enter "STO, 58", then backstep twice until "42" is in the display (for "STO"). He should then overwrite the "STO" command by pressing "Dsz". This will leave "58" in the display. He should then advance the program pointer one step (with "SST") and continue entering the rest of the program.

B. CONTINUOUS K-S PROGRAM LISTING

CARD 1

```
000  76 LBL
001  96 WRT
002  22 INV
003  86 STF
004  03  03
005  53  (
006  53  (
007  43 RCL
008  68  68
009  23 LNX
010  55  ÷
011  02  2
012  23 LNX
013  54  )
014  59 INT
015  32 X↕T
016  02  2
017  45 Y^
018  32 X↕T
019  54  )
020  59 INT
021  42 STO
022  67  67
023  29 CP
024  67  EQ
025  38 SIN
026  87 IFF
027  03  03
028  00  00
029  83  83
030  86 STF
031  03  03
032  25 CLR
033  42 STO
034  65  65
035  42 STO
036  66  66
037  43 RCL
038  67  67
039  44 SUM
040  66  66
041  53  (
042  94 +/-
043  85  +
044  01  1
045  85  +
046  43 RCL
047  68  68
048  54  )
049  42 STO
050  69  69
051  29 CP
052  53  (
053  73 RC*
054  66  66
055  75  -
056  73 RC*
057  65  65
058  54  )
059  77  GE
060  00  00
061  71  71
062  63 EX*
063  66  66
064  63 EX*
065  65  65
066  63 EX*
067  66  66
068  22 INV
069  86 STF
070  03  03
071  01  1
072  44 SUM
073  65  65
074  44 SUM
075  66  66
076  97 DSZ
077  69  69
078  00  00
079  51  51
080  61 GTO
081  00  00
082  26  26
083  22 INV
084  86 STF
085  03  03
086  53  (
087  43 RCL
088  67  67
```

CONTINUOUS K-S, CARD 1 (con't)

```
089  55  ÷
090  02  2
091  54  )
092  61 GTO
093  00  00
094  20  20
095  25 CLR
096  42 STO
097  68  68
098  92 RTN
099  76 LBL
100  15  E
101  25 CLR
102  07  7
103  69 OP
104  17  17
105  86 STF
106  08  08
107  47 CMS
108  92 RTN
109  76 LBL
110  11  A
111  99 PRT
112  72 ST*
113  68  68
114  32 X↑T
115  01  1
116  44 SUM
117  68  68
118  44 SUM
119  59  59
120  43 RCL
121  68  68
122  92 RTN
123  76 LBL
124  13  C
125  71 SBR
126  02  02
127  85  85
128  01  1
129  22 INV
130  44 SUM
131  68  68
132  71 SBR
133  96 WRT
134  98 ADV
135  98 ADV
136  71 SBR
137  02  02
138  60  60
139  98 ADV
140  01  1
141  22 INV
142  90 LST
143  92 RTN
144  76 LBL
145  38 SIN
146  43 RCL
147  59  59
148  42 STO
149  58  58
150  75  -
151  01  1
152  54  )
153  42 STO
154  57  57
155  73 RC*
156  57  57
157  72 ST*
158  58  58
159  01  1
160  94 +/-
161  44 SUM
162  57  57
163  97 DSZ
164  58  58
165  01  01
166  55  55
167  92 RTN
168  76 LBL
169  16 A'
170  43 RCL
171  59  59
172  42 STO
173  58  58
174  73 RC*
175  58  58
176  23 LNX
177  72 ST*
178  58  58
179  97 DSZ
180  58  58
```

CONTINUOUS K-S, CARD 1 (con't)

```
181  01  01
182  74  74
183  18  C'
184  92  RTN
185  76  LBL
186  17  B'
187  43  RCL
188  59  59
189  42  STO
190  58  58
191  73  RC*
192  58  58
193  71  SBR
194  10  E'
195  72  ST*
196  58  58
197  97  DSZ
198  58  58
199  01  01
200  91  91
201  18  C'
202  92  RTN
203  76  LBL
204  18  C'
205  98  ADV
206  98  ADV
207  03  3
208  07  7
209  03  3
210  05  5
211  01  1
212  03  3
213  69  OP
214  01  01
215  03  3
216  01  1
217  03  3
218  06  6
219  02  2
220  01  1
221  03  3
222  02  2
223  03  3
224  05  5
225  69  OP
226  02  02
227  03  3
228  00  0
229  01  1
230  07  7
231  01  1
232  06  6
233  00  0
234  00  0
235  01  1
236  06  6
237  69  OP
238  03  03
239  01  1
240  03  3
241  03  3
242  07  7
243  01  1
244  03  3
245  00  0
246  00  0
247  00  0
248  00  0
249  69  OP
250  04  04
251  69  OP
252  05  05
253  69  OP
254  00  00
255  98  ADV
256  01  1
257  22  INV
258  90  LST
259  92  RTN
260  02  2
261  04  4
262  03  3
263  01  1
264  03  3
265  03  3
266  04  4
267  01  1
268  03  3
269  07  7
270  69  OP
271  02  02
272  01  1
```

CONTINUOUS K-S, CARD 1 (con't)

```
273   06   6
274   01   1
275   03   3
276   03   3
277   07   7
278   01   1
279   03   3
280   69 DP
281   03   03
282   69 DP
283   05   05
284   92 RTN
285   43 RCL
286   59   59
287   75   -
288   01   1
289   54   )
290   42 STO
291   58   58
292   73 RC*
293   58   58
294   32 X:T
295   01   1
296   22 INV
297   44 SUM
298   58   58
299   73 RC*
300   58   58
```

```
301   32 X:T
302   77   GE
303   03   03
304   06   06
305   92 RTN
306   01   1
307   44 SUM
308   58   58
309   97 DSZ
310   58   58
311   02   02
312   92   92
313   71 SBR
314   38 SIN
315   98 ADV
316   98 ADV
317   71 SBR
318   02   02
319   60   60
320   98 ADV
321   01   1
322   22 INV
323   90 LST
324   91 R/S
```

CONTINUOUS K-S, CARD 2

000	76	LBL
001	19	D'
002	77	GE
003	00	00
004	06	06
005	32	X:T
006	92	RTN
007	76	LBL
008	12	B
009	86	STF
010	04	04
011	61	GTO
012	11	A
013	76	LBL
014	13	C
015	86	STF
016	05	05
017	76	LBL
018	11	A
019	25	CLR
020	91	R/S
021	76	LBL
022	16	A'
023	42	STO
024	63	63
025	91	R/S
026	42	STO
027	62	62
028	61	GTO
029	00	00
030	55	55
031	76	LBL
032	17	B'
033	86	STF
034	01	01
035	42	STO
036	63	63
037	01	1
038	42	STO
039	62	62
040	91	R/S
041	42	STO
042	62	62
043	61	GTO
044	00	00
045	55	55
046	76	LBL
047	18	C'
048	86	STF
049	02	02
050	42	STO
051	63	63
052	91	R/S
053	42	STO
054	62	62
055	00	0
056	42	STO
057	68	68
058	42	STO
059	67	67
060	42	STO
061	65	65
062	43	RCL
063	59	59
064	42	STO
065	66	66
066	85	+
067	01	1
068	42	STO
069	64	64
070	54	)
071	42	STO
072	00	00
073	89	π
074	94	+/-
075	72	ST*
076	00	00
077	73	RC*
078	64	64
079	32	X:T
080	01	1
081	44	SUM
082	64	64
083	73	RC*
084	64	64
085	67	EQ
086	01	01
087	46	46
088	01	1

CONTINUOUS K-S, CARD 2 (con't)

089	22	INV
090	44	SUM
091	64	64
092	32	X:T
093	71	SBR
094	01	01
095	83	83
096	42	STO
097	00	00
098	43	RCL
099	64	64
100	55	÷
101	43	RCL
102	59	59
103	54	)
104	75	-
105	43	RCL
106	00	00
107	54	)
108	32	X:T
109	43	RCL
110	68	68
111	19	D'
112	42	STO
113	68	68
114	43	RCL
115	64	64
116	75	-
117	01	1
118	75	-
119	43	RCL
120	65	65
121	54	)
122	55	÷
123	43	RCL
124	59	59
125	54	)
126	94	+/-
127	85	+
128	43	RCL
129	00	00
130	54	)
131	32	X:T
132	43	RCL
133	67	67
134	19	D'
135	42	STO
136	67	67
137	00	0
138	42	STO
139	65	65
140	01	1
141	44	SUM
142	64	64
143	61	GTO
144	01	01
145	49	49
146	01	1
147	44	SUM
148	65	65
149	97	DSZ
150	66	66
151	00	00
152	77	77
153	87	IFF
154	04	04
155	01	01
156	73	73
157	87	IFF
158	05	05
159	01	01
160	78	78
161	43	RCL
162	67	67
163	32	X:T
164	43	RCL
165	68	68
166	19	D'
167	10	E'
168	65	×
169	02	2
170	54	)
171	99	PRT
172	91	R/S
173	43	RCL
174	68	68
175	10	E'
176	99	PRT
177	91	R/S
178	43	RCL
179	67	67
180	10	E'

CONTINUOUS K-S, CARD 2 (con't)

```
181   99 PRT
182   91 R/S
183   87 IFF
184   01  01
185   02  02
186   03  03
187   87 IFF
188   02  02
189   02  02
190   19  19
191   75  -
192   43 RCL
193   63  63
194   54  )
195   55  ÷
196   43 RCL
197   62  62
198   54  )
199   36 PGM
200   19  19
201   12  B
202   92 RTN
203   65  ×
204   43 RCL
205   63  63
206   54  )
207   45 Y^x
208   43 RCL
209   62  62
210   54  )
211   94 +/-
212   22 INV
213   23 LNX
214   94 +/-
215   85  +
216   01  1
217   54  )
218   92 RTN
219   32 X:T
220   43 RCL
221   63  63
222   77  GE
223   02  02
224   48  48
225   43 RCL
226   62  62
227   22 INV
228   77  GE
229   02  02
230   50  50
231   32 X:T
232   75  -
233   43 RCL
234   63  63
235   54  )
236   32 X:T
237   43 RCL
238   62  62
239   75  -
240   43 RCL
241   63  63
242   54  )
243   35 1/X
244   65  ×
245   32 X:T
246   54  )
247   92 RTN
248   00  0
249   92 RTN
250   01  1
251   92 RTN
252   76 LBL
253   10 E'
254   42 STO
255   00  00
256   01  1
257   75  -
258   43 RCL
259   00  00
260   54  )
261   42 STO
262   66  66
263   65  ×
264   43 RCL
265   59  59
266   54  )
267   59 INT
268   85  +
269   01  1
270   54  )
271   42 STO
272   68  68
```

CONTINUOUS K-S, CARD 2 (con't)

```
273  43 RCL
274  59  59
275  42 STO
276  67  67
277  35 1/X
278  42 STO
279  58  58
280  43 RCL
281  00  00
282  42 STO
283  57  57
284  00  0
285  42 STO
286  65  65
287  42 STO
288  61  61
289  01  1
290  94 +/-
291  42 STO
292  56  56
293  71 SBR
294  03  03
295  46  46
296  42 STO
297  60  60
298  43 RCL
299  66  66
300  45 Y^X
301  43 RCL
302  67  67
303  54  )
304  49 PRD
305  60  60
306  43 RCL
307  57  57
308  45 Y^X
309  43 RCL
310  56  56
311  54  )
312  49 PRD
313  60  60
314  43 RCL
315  60  60
316  44 SUM
317  61  61
318  01  1
319  44 SUM
320  65  65
321  44 SUM
322  56  56
323  22 INV
324  44 SUM
325  67  67
326  43 RCL
327  58  58
328  22 INV
329  44 SUM
330  66  66
331  44 SUM
332  57  57
333  97 DSZ
334  68  68
335  02  02
336  93  93
337  43 RCL
338  61  61
339  65  x
340  43 RCL
341  00  00
342  54  )
343  98 ADV
344  98 ADV
345  92 RTN
346  43 RCL
347  65  65
348  42 STO
349  63  63
350  32 X:T
351  00  0
352  22 INV
353  67  EQ
354  03  03
355  58  58
356  01  1
357  92 RTN
358  01  1
359  42 STO
```

```
360   62   62
361   43  RCL
362   63   63
363   94  +/-
364   85    +
365   43  RCL
366   59   59
367   42  STO
368   64   64
369   54    )
370   32  X:T
371   43  RCL
372   63   63
373   22  INV
374   77   GE
375   03   03
376   80   80
377   32  X:T
378   42  STO
379   63   63
380   43  RCL
381   64   64
382   55    ÷
383   43  RCL
384   63   63
385   54    )
386   49  PRD
387   62   62
388   01    1
389   22  INV
390   44  SUM
391   64   64
392   97  DSZ
393   63   63
394   03   03
395   80   80
396   43  RCL
397   62   62
398   92  RTN
```

C. CONTINUOUS K-S EXAMPLE PROBLEMS

1. Example 1

a. Problem:

H_0 : $F(x) = H(x)$ for all x, where H = Uniform $(0,5)$, i.e.,

$$H(x) = \begin{cases} 1 , & x > 5 \\ x/5 , & 0 \leq x \leq 5 \\ 0 , & x < 0 \end{cases}$$

H_1 : $F(x) \neq H(x)$ for some x

Sample size = n = 10

Sample data: 0.8, 4.0, 0.2, 2.6, 3.8, 0.6, 1.0, 4.8, 1.2, 1.4

b. Solution:

Read both sides of Card 1 into the calculator, then:

ENTER	PRESS	DISPLAY	PRINTOUT	COMMENTS
	E	399.69		new partitioning
0.8	A	1	0.8	sample data
4.0	A	2	4.0	
.	.	.	.	
.	.	.	.	
.	.	.	.	
.	.	.	.	
1.4	A	10	1.4	
	C	1	INPUT DATA	[Time: 1 min, 20 sec]
			0.2 01	sorted sample data
			0.6 02	
			.	
			.	
			.	
			4.8 10	
			0. 11	ignore
			0. 12	
			.	
			.	
			.	

(con't)

After completing the previous steps using Card 1 read both sides of Card 2 into the calculator. Do not press "CP" or "RST". Complete the test as follows:

ENTER	PRESS	DISPLAY	PRINTOUT	COMMENTS
	A	Ø		two-tailed test
Ø	C	Ø		H = Uniform (Ø,5)
5	R/S	.2073748895	.2073748895	significance level
				[Time: 1 min, 33 sec]

Since the test conducted was a two-tailed test the significance level computed is approximate. The user is assured that the true significance level is no greater than the value computed. The user may recall the test statistic t = .32 from register ØØ.

2. Example 2

a. Problem:

H_o : $F(x) \leq H(x)$ for all x, where H = Normal $(3, 2^2)$

H_1 : $F(x) > H(x)$ for some x

Sample size = n = 5

Sample data: 1.462, -Ø.311, Ø.555, 5.711, -Ø.Ø78

b. Solution:

Read both sides of card 1 into the calculator, then:

ENTER	PRESS	DISPLAY	PRINTOUT	COMMENTS
	E	399.69		new partitioning
1.462	A	1	1.462	sample data
-Ø.311	A	2	-Ø.311	
Ø.555	A	3	Ø.555	
5.711	A	4	5.711	
-Ø.Ø78	A	5	-Ø.Ø78	
	C	1	INPUT DATA	[Time: 3Ø sec]
			-Ø.311 Ø1	Sorted Sample Data
			-Ø.Ø78 Ø2	
			Ø.555 Ø3	
			1.462 Ø4	
			5.711 Ø5	
			Ø. Ø6	ignore
			Ø. Ø7	
			.	
			.	
			.	

Read Card 2, both sides, into the calculator, then:

ENTER	PRESS	DISPLAY	PRINTOUT	COMMENTS
	B	Ø	.	one-tailed test
3	A′	3		H = Normal $(3, 2^2)$
2	R/S	.Ø2Ø1689866	.Ø2Ø1689866	significance level [Time: 53 sec]

The user may recall the test statistic t = .5790534154 from register ØØ.

3. Example 3

a. Problem:

H_0: $F(x) = H(x)$ for all x, where

$$H(x) = \begin{cases} 0 , & x < 0 \\ x^2/4 , & 0 \leq x \leq 2 \\ 1 , & x > 2 \end{cases}$$

H_1: $F(x) \neq H(x)$ for some x

Sample size n = 10

Sample data: 1.8, 1.0, 1.3, 0.5, 1.98, 0.95, 1.91, 0.75, 1.85, 1.6

b. Discussion:

Continuous K-S does not have a subroutine to evaluate the hypothesized CDF H for this problem. This is an ideal situation to invoke the probability integral transformation. Under the null hypothesis X has CDF H, thus, H(X) is distributed uniform (0,1) under the null hypothesis. If we transorm each sample data value with H, defined above, then we can proceed with the K-S test using a uniform (0,1) distribution in place of the orginal hypothesized distribution.

c. Solution:

Read both sides of Card 1 into the calculator, then:

ENTER	PRESS	DISPLAY	PRINTOUT	COMMENT
	E	399.69		new partitioning
1.8	A	1	1.8	
1.0	A	2	1.0	
.	.	.	.	sample data
.	.	.	.	
.	.	.	.	
1.6	A	10	1.6	
	C	1	INPUT DATA	[Time: 1 min, 20 sec]
			0.5 01	
			0.75 02	
			.	sorted sample data
			.	
			.	
			1.98 10	
			0. 11	
			0. 12	
			.	ignore
			.	
			.	
	GTO 325	1		
	LRN	325 00		
	LBL	326 00		
	E′	327 00		
	x^2	328 00		Transformation
	÷	329 00		Subroutine H
	4	330 00		
	)	331 00		
	INV SBR	332 00		
	LRN	1		
2	E′	1		Subroutine test
	B′	TRANSFORMED DATA		
		0.0625 01		
		0.140625 02		
		0.225625 03		
		0.25 04		
		0.4225 05		Transformed data
		0.64 06		
		0.81 07		
		0.855625 08		
		0.912025 09		
		0.9801 10		
		0. 11		
		0. 12		ignore
		.		
		.		

Read both sides of Card 2 into the calculator, then:

ENTER	PRESS	DISPLAY	PRINTOUT	COMMENT
	A	Ø		two-sided test
Ø	C′	Ø		uniform (Ø,1)
1	R/S	.7275048241	.7275048421	upper bound on significance level
				[Time: 2 min]

If the significance level computed had been .1 or less we could consider it exact. With as value as large as that computed, however, we should consider it an upper bound on the true significance level and conclude that the sample data and the null hypothesis are in close agreement. The user may recall the test statistic t = .21 from register ØØ.

4. Example 4

a. Problem:

H_0 : $F(x) \geq H(x)$ for all x, where H = lognormal $(3, 5^2)$

H_1 : $F(x) < H(x)$ for some x

Sample size = n = 7

Sample data: 99.31, 22.47, .0608, 3294.5, 4.95, 14.88, 29.96

b. Discussion:

This problem demonstrates the use of the natural log transformation. It is appropriate when the hypothesized distribution is lognormal. Recall that X is distributed lognormal (μ, σ^2) if and only if ln(X) is distributed normal (μ, σ^2). We rely on the natural log transformation subroutine incorporated into Card 1 of Continuous K-S to transform our sample data. We then test the transformed data against a normal $(3, 5^2)$ CDF using Card 2.

c. Solution:

Read both sides of Card 1 into the calculator, then:

ENTER	PRESS	DISPLAY	PRINTOUT	COMMENT
	E	399.69		new partitioning
99.31	A	1	99.31	
22.47	A	2	22.47	
0.0608	A	3	0.0608	
3294.5	A	4	3294.5	sample data
4.95	A	5	4.95	
14.88	A	6	14.88	
29.96	A	7	29.96	

ENTER	PRESS	DISPLAY	PRINTOUT		COMMENT
	C	1	INPUT DATA		[Time: 46 sec]
			0.0608	01	sorted sample data
			4.95	02	
			14.88	03	
			22.47	04	
			29.96	05	
			99.31	06	
			3294.5	07	
			0.	08	ignore
			0.	09	
			.		
			.		
			.		
	A′	1	TRANSFORMED DATA		
			-2.80016549	01	Transformed data
			1.599387577	02	
			2.700018029	03	
			3.112181086	04	
			3.399863159	05	
			4.598246271	06	
			8.10000969	07	
			0.	08	ignore
			0.	09	
			.		
			.		
			.		

Read both sides of Card 2 into the calculator, then:

ENTER	PRESS	DISPLAY	PRINTOUT	COMMENT
	C	0		One-sided test
3	A′	3		normal $(3, 5^2)$
5	R/S	.3672246357	.3672246357	significance level
				[Time 1 min, 25 sec]

The user may recall test statistic t = .2468346856 from register 00.

APPENDIX B

DISCRETE K-S: ENTRY PROCEDURE, LISTING, AND EXAMPLES

A. ENTRY PROCEDURE

The program is organized in two distinct parts, Card 1 and Card 2. Card 1 should be recorded on a magnetic card using banks 1 and 2 [Ref. 8: pp. VII-2] with the calculator in standard partitioning [Ref. 8: pp. V-22]. Card 2 should also be recorded in banks 1 and 2 but with partitioning 399.69.

It should be noted in the program lising below that Cards 1 and 2 both use the "Dsz" command followed by register numbers greater than 9. The user's manual [Ref. 8: pp. V-63] implies that this will not work but, in fact, it does. The user has to use a special entry technique to get the program steps into the calculator correctly, however. One very convenient way to accomplish this is illustrated easily by an example.

Assume the steps at one point in the program are "Dsz, 58". The user can first enter "STO, 58", then backstep twice until "42" is in the display (for "STO"). He should then overwrite the "STO" command by pressing "Dsz". This will leave "58" in the display. He should advance the program pointer one step (with "SST") and continue entering the rest of the program.

Discrete K-S uses the "HIR" command, explained in detail in Appendix D. It is represented by machine code "82" and thus must be entered with a special procedure. A convenient way to enter "HIR" is to precede it with "STO", enter "82", and backstep twice, leaving "42" (for "STO") in the display. Delete the "STO" command with "del", advance the program pointer one step, and continue entering the program. When "HIR" is followed by another two-digit merged code, e.g., "HIR, 16", the entry procedure should be used twice, the first time to enter "82" and the second time to enter "16".

B. DISCRETE K-S PROGRAM LISTING

CARD 1

```
000  76 LBL
001  15  E
002  25 CLR
003  47 CMS
004  86 STF
005  00  00
006  61 GTO
007  00  00
008  16  16
009  76 LBL
010  10 E'
011  22 INV
012  86 STF
013  00  00
014  25 CLR
015  47 CMS
016  22 INV
017  86 STF
018  04  04
019  91 R/S
020  42 STO
021  41  41
022  85  +
023  01  1
024  54  )
025  32 X:T
026  43 RCL
027  41  41
028  91 R/S
029  42 STO
030  42  42
031  01  1
032  42 STO
033  00  00
034  91 R/S
035  72 ST*
036  00  00
037  22 INV
038  87 IFF
039  00  00
040  00  00
041  44  44
042  44 SUM
043  45  45
044  69 OP
045  20  20
046  43 RCL
047  00  00
048  22 INV
049  77  GE
050  00  00
051  34  34
052  22 INV
053  87 IFF
054  00  00
055  16 A'
056  01  1
057  93  .
058  00  0
059  01  1
060  32 X:T
061  43 RCL
062  45  45
063  77  GE
064  03  03
065  87  87
066  32 X:T
067  93  .
068  09  9
069  09  9
070  77  GE
071  03  03
072  87  87
073  76 LBL
074  16 A'
075  43 RCL
076  41  41
077  85  +
078  01  1
079  54  )
080  32 X:T
081  02  2
082  01  1
083  42 STO
084  00  00
085  01  1
086  42 STO
087  44  44
088  91 R/S
```

DISCRETE K-S, CARD 1 (con't)

```
089  72 ST*
090  00  00
091  44 SUM
092  43  43
093  69 DP
094  20  20
095  01  1
096  44 SUM
097  44  44
098  43 RCL
099  44  44
100  22 INV
101  77  GE
102  00  00
103  88  88
104  43 RCL
105  42  42
106  32 X:T
107  43 RCL
108  43  43
109  22 INV
110  67  EQ
111  03  03
112  91  91
113  25 CLR
114  91 R/S
115  76 LBL
116  11  A
117  86 STF
118  01  01
119  22 INV
120  86 STF
121  02  02
122  61 GTO
123  01  01
124  42  42
125  76 LBL
126  12  B
127  22 INV
128  86 STF
129  02  02
130  22 INV
131  86 STF
132  01  01
133  61 GTO
134  01  01
135  42  42
136  76 LBL
137  13  C
138  86 STF
139  01  01
140  86 STF
141  02  02
142  87 IFF
143  04  04
144  03  03
145  01  01
146  86 STF
147  04  04
148  22 INV
149  87 IFF
150  00  00
151  01  01
152  70  70
153  43 RCL
154  45  45
155  35 1/X
156  42 STO
157  45  45
158  43 RCL
159  41  41
160  42 STO
161  44  44
162  43 RCL
163  45  45
164  64 PD*
165  44  44
166  97 DSZ
167  44  44
168  01  01
169  64  64
170  43 RCL
171  42  42
172  35 1/X
173  42 STO
174  43  43
175  02  2
176  01  1
177  42 STO
178  00  00
179  43 RCL
180  41  41
```

DISCRETE X-S, CARD 1 (con't)

181	42	STO	227	43	43
182	44	44	228	97	DSZ
183	43	RCL	229	44	44
184	43	43	230	02	02
185	64	PI*	231	08	08
186	00	00	232	01	1
187	01	1	233	42	STO
188	44	SUM	234	00	00
189	00	00	235	42	STO
190	97	DSZ	236	50	50
191	44	44	237	42	STO
192	01	01	238	51	51
193	83	83	239	02	2
194	43	RCL	240	01	1
195	41	41	241	42	STO
196	75	-	242	43	43
197	01	1	243	43	RCL
198	54	)	244	41	41
199	42	STO	245	42	STO
200	44	44	246	44	44
201	02	2	247	00	0
202	01	1	248	42	STO
203	42	STO	249	47	47
204	00	00	250	42	STO
205	01	1	251	48	48
206	42	STO	252	73	RC*
207	43	43	253	00	00
208	73	RC*	254	75	-
209	00	00	255	73	RC*
210	69	OP	256	43	43
211	20	20	257	54	)
212	74	SM*	258	42	STO
213	00	00	259	45	45
214	22	INV	260	94	+/-
215	87	IFF	261	32	X:T
216	00	00	262	43	RCL
217	02	02	263	48	48
218	28	28	264	77	GE
219	73	RC*	265	02	02
220	43	43	266	77	77
221	32	X:T	267	32	X:T
222	01	1	268	42	STO
223	44	SUM	269	48	48
224	43	43	270	43	RCL
225	32	X:T	271	00	00
226	74	SM*	272	42	STO

```
273   50   50
274   61  GTO
275   02   02
276   92   92
277   43  RCL
278   45   45
279   32  X:T
280   43  RCL
281   47   47
282   77   GE
283   02   02
284   92   92
285   32  X:T
286   42  STO
287   47   47
288   43  RCL
289   00   00
290   42  STO
291   51   51
292   01   1
293   44  SUM
294   00   00
295   44  SUM
296   43   43
297   97  DSZ
298   44   44
299   02   02
300   52   52
301   87  IFF
302   01   01
303   03   03
304   12   12
305   43  RCL
306   47   47
307   82  HIR
308   04   04
309   61  GTO
310   03   03
311   31   31
312   87  IFF
313   02   02
314   03   03
315   27   27
316   43  RCL
317   47   47
318   32  X:T
319   43  RCL
320   48   48
321   77   GE
322   03   03
323   27   27
324   61  GTO
325   03   03
326   05   05
327   43  RCL
328   48   48
329   82  HIR
330   04   04
331   07   7
332   69  OP
333   17   17
334   01   1
335   75   -
336   82  HIR
337   14   14
338   95   =
339   65   ×
340   43  RCL
341   42   42
342   95   =
343   85   +
344   02   2
345   95   =
346   59  INT
347   42  STO
348   59   59
349   43  RCL
350   42   42
351   82  HIR
352   02   02
353   43  RCL
354   41   41
355   82  HIR
356   03   03
357   05   5
358   05   5
359   75   -
```

DISCRETE K-S, CARD 1 (con't)

360	82	HIR
361	13	13
362	95	=
363	32	X:T
364	43	RCL
365	59	59
366	77	GE
367	03	03
368	84	84
369	93	.
370	00	0
371	00	0
372	00	0
373	00	0
374	00	0
375	00	0
376	00	0
377	00	0
378	01	1
379	42	STO
380	56	56
381	82	HIR
382	14	14
383	91	R/S
384	04	4
385	04	4
386	91	R/S
387	32	X:T
388	02	2
389	02	2
390	91	R/S
391	32	X:T
392	03	3
393	03	3
394	91	R/S

DISCRETE K-S, CARD 2

```
000  76 LBL
001  12  B
002  82 HIR
003  13  13
004  75  -
005  82 HIR
006  12  12
007  95  =
008  32 X;T
009  82 HIR
010  13  13
011  85  +
012  82 HIR
013  12  12
014  95  =
015  91 R/S
016  76 LBL
017  13  C
018  82 HIR
019  12  12
020  91 R/S
021  76 LBL
022  14  D
023  82 HIR
024  13  13
025  91 R/S
026  76 LBL
027  11  A
028  22 INV
029  87 IFF
030  01  01
031  00  00
032  81  81
033  22 INV
034  87 IFF
035  02  02
036  00  00
037  43  43
038  71 SBR
039  88 DMS
040  82 HIR
041  06  06
042  91 R/S
043  71 SBR
044  88 DMS
045  82 HIR
046  06  06
047  22 INV
048  86 STF
049  01  01
050  71 SBR
051  88 DMS
052  82 HIR
053  05  05
054  85  +
055  82 HIR
056  16  16
057  95  =
058  42 STO
059  00  00
060  82 HIR
061  15  15
062  65  ×
063  82 HIR
064  16  16
065  55  ÷
066  02  2
067  95  =
068  82 HIR
069  02  02
070  32 X;T
071  82 HIR
072  12  12
073  94 +/-
074  85  +
075  43 RCL
076  00  00
077  95  =
078  82 HIR
079  03  03
080  91 R/S
081  71 SBR
082  88 DMS
083  82 HIR
084  05  05
085  91 R/S
086  76 LBL
087  88 DMS
088  82 HIR
```

DISCRETE K-S, CARD 2 (con't)

```
099  13  13
090  85  +
091  01  1
092  95  =
093  42  STO
094  62  62
095  42  STO
096  58  58
097  01  1
098  72  ST*
099  62  62
100  02  2
101  42  STO
102  00  00
103  01  1
104  82  HIR
105  07  07
106  00  0
107  42  STO
108  60  60
109  01  1
110  42  STO
111  69  69
112  42  STO
113  64  64
114  43  RCL
115  59  59
116  42  STO
117  68  68
118  43  RCL
119  69  69
120  71  SBR
121  03  03
122  05  05
123  32  X:T
124  00  0
125  67  EQ
126  01  01
127  55  55
128  32  X:T
129  45  Y^x
130  43  RCL
131  64  64
132  95  =
133  42  STO
134  57  57
135  43  RCL
136  69  69
137  75  -
138  01  1
139  95  =
140  82  HIR
141  08  08
142  71  SBR
143  02  02
144  68  68
145  49  PRD
146  57  57
147  73  RC*
148  62  62
149  49  PRD
150  57  57
151  43  RCL
152  57  57
153  44  SUM
154  60  60
155  01  1
156  44  SUM
157  69  69
158  44  SUM
159  62  62
160  94  +/-
161  44  SUM
162  64  64
163  43  RCL
164  00  00
165  32  X:T
166  43  RCL
167  69  69
168  22  INV
169  77  GE
170  01  01
171  18  18
172  01  1
173  75  -
174  43  RCL
175  60  60
176  95  =
177  72  ST*
178  62  62
179  01  1
180  42  STO
```

DISCRETE K-S, CARD 2 (con't)

```
181   69   69
182   44  SUM
183   00   00
184   82  HIR
185   37   37
186   82  HIR
187   17   17
188   42  STO
189   64   64
190   43  RCL
191   68   68
192   32  X;T
193   43  RCL
194   00   00
195   77   GE
196   02   02
197   08   08
198   00   0
199   42  STO
200   60   60
201   43  RCL
202   58   58
203   42  STO
204   62   62
205   61  GTO
206   01   01
207   18   18
208   01   1
209   94  +/-
210   44  SUM
211   68   68
212   43  RCL
213   58   58
214   42  STO
215   62   62
216   00   0
217   82  HIR
218   08   08
219   42  STO
220   61   61
221   82  HIR
222   12   12
223   82  HIR
224   07   07
225   42  STO
226   67   67
227   71  SBR
228   02   02
229   68   68
230   42  STO
231   60   60
232   73  RC*
233   62   62
234   49  PRD
235   60   60
236   43  RCL
237   69   69
238   71  SBR
239   03   03
240   05   05
241   45  Y^
242   43  RCL
243   67   67
244   95   =
245   49  PRD
246   60   60
247   43  RCL
248   60   60
249   44  SUM
250   61   61
251   01   1
252   44  SUM
253   62   62
254   44  SUM
255   69   69
256   82  HIR
257   38   38
258   94  +/-
259   44  SUM
260   67   67
261   97  DSZ
262   68   68
263   02   02
264   27   27
265   43  RCL
266   61   61
267   92  RTN
268   00   0
269   32  X;T
270   82  HIR
271   18   18
272   67   EQ
```

DISCRETE K-S, CARD 2 (con't)

```
273   08   08
274   03   03
275   42 STO
276   66   66
277   82 HIR
278   17   17
279   42 STO
280   65   65
281   01   1
282   42 STO
283   63   63
284   43 RCL
285   65   65
286   55   ÷
287   43 RCL
288   66   66
289   95   =
290   49 PRD
291   63   63
292   01   1
293   94 +/-
294   44 SUM
295   65   65
296   97 DSZ
297   66   66
298   02   02
299   84   84
300   43 RCL
301   63   63
302   92 RTN
303   01   1
304   92 RTN
305   22 INV
306   87 IFF
307   01   01
308   03   03
309   64   64
310   94 +/-
311   85   +
312   82 HIR
313   12   12
314   85   +
315   01   1
316   95   =
317   55   ÷
318   82 HIR
319   12   12
320   95   =
321   75   -
322   82 HIR
323   14   14
324   95   =
325   85   +
326   43 RCL
327   56   56
328   95   =
329   32 X:T
330   43 RCL
331   01   01
332   77   GE
333   03   03
334   91   91
335   82 HIR
336   13   13
337   42 STO
338   63   63
339   73 RC*
340   63   63
341   22 INV
342   77   GE
343   03   03
344   49   49
345   97 DSZ
346   63   63
347   03   03
348   39   39
349   22 INV
350   87 IFF
351   01   01
352   03   03
353   55   55
354   92 RTN
355   01   1
356   44 SUM
357   63   63
358   01   1
359   75   -
```

```
360   73 RC*
361   63  63
362   95  =
363   92 RTN
364   75  -
365   01  1
366   95  =
367   55  ÷
368   82 HIR
369   12  12
370   95  =
371   85  +
372   82 HIR
373   14  14
374   75  -
375   43 RCL
376   56  56
377   95  =
378   32 X↕T
379   82 HIR
380   13  13
381   75  -
382   01  1
383   95  =
384   42 STO
385   63  63
386   73 RC*
387   63  63
388   77  GE
389   03  03
390   35  35
391   00  0
392   92 RTN
393   76 LBL
394   15  E
395   06  6
396   69 OP
397   17  17
398   91 R/S
```

C. DISCRETE K-S EXAMPLE PROBLEMS

1. Example 1

a. Problem:

H_o : $F(x) = H(x)$ for all x, where H is defined below

H_1 : $F(x) \neq H(x)$ for some x

Number of cells = m = 5

Sample size = n = 10

	CELL NUMBER (i)				
	1	2	3	4	5
mass value (h_i):	.2	.2	.2	.2	.2
cumulative mass value (H_i):	.2	.4	.6	.8	1.0
Number of observations (X_i):	3	3	4	0	0

b. Solution:

Read both sides of Card 1 into the calculator, then:

COMMENT	ENTER	PRESS	DISPLAY	COMMENT
initialize		E	0	H defined with h_i s
no. of cells	5	R/S	5	
sample size	10	R/S	1	
cell mass values	.2	R/S	2	cell number of next mass value entry
	.2	R/S	3	
	.2	R/S	4	
	.2	R/S	5	
	.2	R/S	1	
observations per cell (frequency count)	3	R/S	2	cell number of next frequency count
	3	R/S	3	
	4	R/S	4	
	0	R/S	5	
	0	R/S	0	
		A	.4	test statistic of two-sided test [Time: 22 sec]

Read both sides of Card 2 into the calculator, then:

PRESS	DISPLAY	COMMENT
A	.0414007042	point estimate of significance level [Time: 7 min]
B	.0416172032	upper bound on significance level
x:t	.0411842053	lower bound on significance level
C	.000216499	maximum error in point estimate
D	.0414007042	recall point estimate
E	479.59	return TI-59 to normal partitioning

The user can perform the same K-S test by initializing Card 1 with "E′" and then entering H_i values instead of h_i values. The test results will be the same.

2. Example 2

a. Problem:

H_o : $F(x) \geq H(x)$ for all x, where H is defined below

H_1 : $F(x) < H(x)$ for some x

Number of cells = m = 3

Sample size = n = 15

	CELL NUMBER (i) 1	2	3
mass value (h_i):	.3624	.4167	.2209
cumulative mass value (H_i)	.3624	.7791	1.0
number of observations (X_i)	5	3	7

b. Solution:

Read both sides of Card 1 into the calculator, then:

COMMENT	ENTER	PRESS	DISPLAY	COMMENT
initialize		E′	Ø	H defined with H_i s
no.of cells	3	R/S	3	
sample size	15	R/S	1	cell number of next
cumulative	.3624	R/S	2	cumulative mass
cell mass	.7791	R/S	3	value
values	1.Ø	R/S	1	
				cell number of next
observations	5	R/S	2	frequency count
per cell	3	R/S	3	
(frequency	7	R/S	Ø	
counts)				
		B	.245766667	test statistic of one-sided test [Time: 15 sec]

Read both sides of Card 2 into the calculator, then:

PRESS	DISPLAY	COMMENT
A	.0395671995	significance level [Time: 11 min]
E	479.59	return TI-59 to normal partitioning

APPENDIX C

CONDENSED USER INSTRUCTIONS

The following two pages are condensed user instructions for Continuous K-S and Discrete K-S, respectively. It is recommended that the reader familiarize himself with the more expanded discussion of programs in sections IV and V above before relying on the condensed instructions.

The condensed instructions for both programs are written for a complete K-S test. It is possible to use Continuous K-S to evaluate a known test statistic from a continuous K-S test. The procedure is as follows.

STEP	PROCEDURE	ENTER	PRESS	DISPLAY
1.	Repartition Ti-59	7	Op 17	399.69
2.	Enter both sides of Card 2 (Continuous K-S)			
3.	enter sample size	n	STO 59	n
4.	enter test statistic	t	STO ØØ	t
5.	compute significance level		E′	S.L.

If the test was one-sided the significance level computed is exact. If the test was two-sided the result should be doubled. The doubled value is an upper bound on the exact significance level.

CONTINUOUS K-S

Card 1 (banks 1 & 2): standard partitioning — Applied Statistics Module
Card 2 (banks 1 & 2): 399.69 partitioning — Printer optional

Purpose: K-S test on continuous random variable with unknown CDF, F. Maximum sample size n = 54. Direction of test specified by null hypothesis (H_o). Alternative hypothesis (H_1) is negation of H_o. Hypothesized CDF, H, can be uniform, normal, Weibull, or exponential. If another H is desired the user can add a subroutine transforming the desired H to one of the above distributions. Program computes exact significance level (S.L.) for a one-sided test and an upper bound on S.L. for two-sided test.

USER INSTRUCTIONS

STEP	PROCEDURE	ENTER	PRESS	DISPLAY	PRINTOUT
1	Enter both sides of Card 1				
2.	Initialize		E	399.69	
3.	Enter sample data	X_i ($i=1, 2, \ldots, n$)	A	i	X_i
4.	Sort data		C	1	$X_{(1)}, \ldots, X_{(n)}$

5. Tranforming data (optional step; normally omitted): If hypothesized CDF, H, is NOT uniform, normal, Weibull, or exponential then the user may enter a subroutine to transformed H to one of those distributions. The subroutine must begin with "LBL E′" and end with "INV SBR". For a transformation M the subroutine should be written to tranform some value c to M(c) with c in the x-register when the subroutine is called and M(c) is returned to the main program. Use only registers n+1 through 57 (sample size = n) and limit subroutine to 75 steps. After subroutine entered and validated, press "B′". $M(X_{(n)})$ will be displayed.
   If user wishes to transform each sample value x to ln(x) press "A′". The transformation subroutine is already provided and is useful to transform lognormal (μ,σ^2) data to normal (μ,σ^2). Do not press "B′".
   Proceed with test with the new distribution substituted for original H.

STEP	PROCEDURE	ENTER	PRESS	DISPLAY	PRINTOUT
6.	Enter both sides of Card 2				
7.	Choose direction of test:				
	If H_o: $F(x) = H(x)$		A	Ø	
	If H_o: $F(x) \geq H(x)$		B	Ø	
	If H_o: $F(x) \leq H(x)$		C	Ø	
8.	Specify H for test:				
	If H = normal (μ,σ^2)	μ	A′	μ	
		σ	R/S	S.L.	S.L.
	If H = exponential (λ)	λ	B′	1	
		1	R/S	S.L.	S.L.
	If H = Weibull (λ, α)	λ	B′	1	
		α	R/S	S.L.	S.L.
	If H = uniform (a, b)	a	C′	a	
		b	R/S	S.L.	S.L.
9.	The test statistic can be recalled from register Ø0.				

DISCRETE K-S

Card 1 (banks 1 & 2): standard partitioning — Any Library Module
Card 2 (banks 1 & 2): 399.69 partitioning — Does not use printer

Purpose: K-S test on discrete random variable with unknown CDF, F. For sample size n and random variable with m possible outcomes (cells) the program will handle any problem with $n+m \leq 54$. Maximum number of cells: 20. If n+m slightly exceeds 54 program will attempt to solve; edit code "44" displayed if not possible. Program computes significance level of test.

Notation:
- X_i = number of observations of X in cell i; i = 1, . . . , m
- h_i = P(random observation is in cell i)
- H_i = P(random observation is in cell j; $j \leq i$)
- t = test statistic
- H_o = null hypothesis
- H = hypothesized CDF of random variable X
- F = actual CDF of random variable X (unknown)

The hypothesized CDF, H, may be defined by specifying h_i or H_i values.

USER INSTRUCTIONS

STEP	PROCEDURE	ENTER	PRESS	DISPLAY
1.	Initialize:			
	If H defined by h_i values		E	0
	If H defined by H_i values		E′	0
2.	Specify number of cells	m	R/S	m
3.	Specify sample size	n	R/S	1
4.	Define H with either h_i values or H_i values	h_i or H_i ($i = 1, \ldots, m-1$)	R/S	i+1
		h_m or H_m	R/S	1 (note 1)
2.	Specify cell frequency counts	X_i ($i = 1, \ldots, m-1$)	R/S	i+1
		X_m	R/S	0 (note 2)
5.	Specify direction of test			
	If H_o: $F(x) = H(x)$		A	t
	If H_o: $F(x) \geq H(x)$		B	t
	If H_o: $F(x) \leq H(x)$		C	t (note 3)
6.	Enter both sides of Card 2			
7.	Compute significance level (S.L.)		A	S.L.
8.	Evaluate S.L. for two-sided test; (S.L. exact for one-sided test)		B	max S.L.
			x≷t	min S.L.
			C	max error
			D	recall S.L.

Note 1: Edit code "22" indicates ($h_1 + \ldots + h_m$) greater than 1.01 or less than .99. User may press "A" to normalize h_i's & continue.
Note 2: Edit code "33" indicates ($X_1 + \ldots + X_m$) ≠ n. Cannot proceed.
Note 3: Edit code "44" indicates problem too large. Recall t from Reg 00.

APPENDIX D

USE OF PENDING ARITHMETIC REGISTERS IN THE TI-59

The TI-59 calculator has eight pending arithmetic registers. They are not accessible from the keyboard but are accessible from within a program and are used extensively in Discrete K-S. A detailed discussion of these special registers has been provided by a non-profit organization called the SR-52 Users Club [Ref. 10]. Unfortunately the user's manual provided by the manufacturer [Ref. 8] does not discuss them.

Each of the eight pending registers can be accessed with "82" followed by a two digit number, represented here as "mn". The access code "82" is represented as "HIR" by a Texas Instruments printer. This mnemonic has been interpreted as both "Hierarchy Internal Register" and "Hidden Internal Register".

The value of m determines which register operation is performed, i.e.,

m	operation
0	STO
1	RCL
3	SUM
4	Prd
5	INV SUM
6, 7, 8, or 9	INV Prd .

The value of n (1, 2, . . . , or 8) determines which HIR is

accessed. For example, to add a number directly to the value currently in HIR 4 the program codes "82, 34" would be used.

To enter a two digit code like "82" or "34" in a program it is convenient to "trick" the calulator with a register command. For instance, with the calculator in the "LRN" mode press "STO, 8, 2" then backstep (with "BST") twice and "42" (for "STO") will be displayed. Delete the register command "STO" by pressing "del" and the "82" remains as a merged code. Advance the program pointer with "SST" one step and the remaining program steps can be entered.

The pending arithmetic registers serve several purposes. When nested arithmetic operations are performed the operands are first pushed into HIR 1 and then into 2, 3, . . ., 8. Use of the Op codes 1, 2, 3, and 4 assign values to HIRs 5, 6, 7, and 8, respectively. Other keyboard operations which affect the HIRs are "P→R" and "D.MS" .

After extensive experimentation with HIRs the writer is convinced that the TI-59 uses HIRs in other ways besides those mentioned. For instance, Discrete K-S, Card 2, does not use nested arithmetic operations, "P→R", or "D.MS". It uses HIRs 2 - 7. When the program was modified to also use HIR 1 it did not perform computations correctly. For some reason the manufacturer has not disclosed any information about the HIRs except to acknowledge their existence.

Utilizing HIRs modestly increases the calculator's data storage capacity. It also provides a place to store a few

numbers to be retained after "CM" is pressed. When incorporating HIRs into a program the following approach is recommended. First, write the program without using any HIRs and validate it. Second, incorprate the desired HIRs to free regular data registers and then revalidate the program. Trying to find an error in a program that was written with HIRs initially can be very frustrating; the program may be affected by some mysterious HIR-related problem that the user may never find.

LIST OF REFERENCES

1. Dixon, W.J. and Massey, F.J., Introduction to Statistical Analysis, 3d ed., McGraw-Hill, 1969.

2. Kolmogorov, A., "Sulla Determinazione Empirica di una Legge di Distribuzione," Giornale dell' Istituto Italiano degle Attuare, v. 4, p. 83-91, 1933.

3. Smirnov, N., "Table for Estimating the Goodness-of-Fit of Empirical Distributions," Annals of Mathematical Statistics, v. 19, p. 279, 1948.

4. Bradley, J.V., Distribution-Free Statistical Analysis, Prentice Hall, 1968.

5. Conover, W.J., "A Kolmogorov Goodness-of-Fit Test for Discontinuous Distributions," Journal of the American Statistical Association, v. 67, no. 339, p.591, September 1972.

6. Mood, A.M., Graybill, F.A., and Boes, D.C., Introduction to the Theory of Statistics, 3d ed., McGraw-Hill, 1974.

7. Birnbaum, Z.W. and Tingey, F.H., "One-Sided Confidence Contours for Probability Distribution Functions," Annals of Mathematical Statistics, v. 22, p. 592, 1951.

8. Texas Instuments, Personal Programming, 1977.

9. Allen, M.E., Kolmogorov-Smirnov Tests for Discrete Distributions, M.S. Thesis, Naval Postgraduate School, 1976.

10. SR-52 Users Club, 52-Notes, v. 2, no. 9, September 1977.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, California 93940	2
3.	Department Chairman, Code 55 Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
4.	Professor Donald R. Barr Naval Postgraduate School, Code 55Bn Monterey, California 93940	1
5.	Professor Gary K. Poock Naval Postgraduate School, Code 55Pk Monterey, California 93940	1
7.	Professor Peter C.C. Wang Naval Postgraduate School, Code 53Wg Monterey, California 93940	1
8.	LT Elena W. Brown PATWING ONE DET CUBI Box 54 FPO San Francisco, California 96654	1
9.	CPT David F. Davis Department of Mathematics United States Military Academy Attn: MADN-A West Point, New York 10996	1
10.	LT Douglas W. Brown NAS Cubi Point, Box 21 FPO San Francisco, California 96654	1

Thesis
B80996
c.1

192311

Brown
The one-sample K-S test: TI-59 programs for discrete and continuous random variables.

Thesis
B80996
c.1

192311

Brown
The one-sample K-S test: TI-59 programs for discrete and continuous random varaibles.

thesB80996
The one-sample K-S test :
3 2768 001 01876 5
DUDLEY KNOX LIBRARY